Improving Neutrino Energy Reconstruction with Recurrent Neural Networks at NOvA

Dmitrii Torbunov, University of Minnesota
torbu001@umn.edu

February 25, 2021

Introduction

- In this presentation, I will discuss development of the RNN neutrino energy estimator for the NOvA neutrino oscillation experiment.

- NOvA (NuMI Off-Axis ν_e Appearance) – a long baseline accelerator based neutrino oscillation experiment.

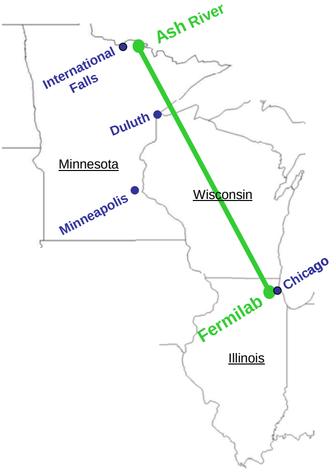
- Plan of the talk:
  - Overview of the NOvA experiment.
  - Overview of the neutrino energy estimation at NOvA.
  - Development of the RNN energy estimator.
  - Other applications of the RNN architecture.

Neutrino Oscillation

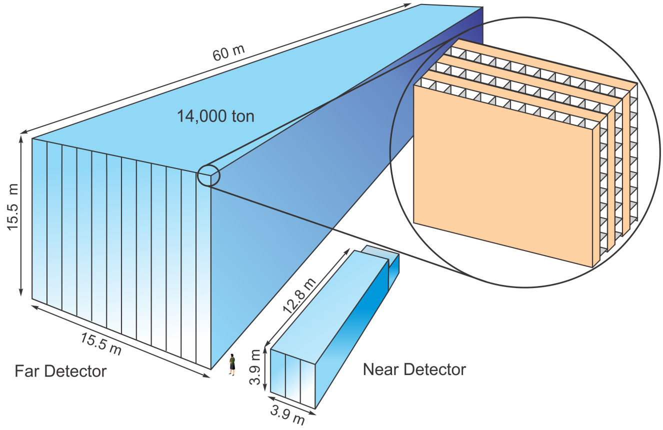
- Three generations (flavors) of neutrinos are known: ν_e, ν_μ, ν_τ.
- It was discovered, that neutrinos change their flavor over time.
- Probability $P_{\nu_\alpha \to \nu_\beta}$ of neutrino changing its flavor is a periodic function of time – phenomenon known as Neutrino Oscillation.
- By measuring neutrino oscillation probability $P_{\nu_\alpha \to \nu_\beta}$ we can get estimates of the fundamental parameters of the neutrino physics: Δm^2_{21}, Δm^2_{32}, θ_{12}, θ_{23}, θ_{13}, δ_{CP}

NOvA Overview

- NOvA is a long-baseline (810 km) accelerator based neutrino oscillation experiment.
- Studies NuMI muon (anti-) neutrino beam (700 kW) produced at Fermilab.
- NOvA detects neutrinos with two finely grained liquid scintillator detectors.



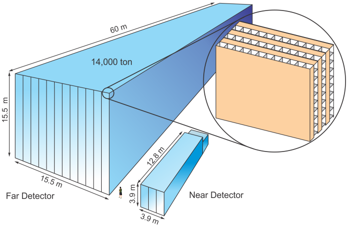
NOvA Detectors

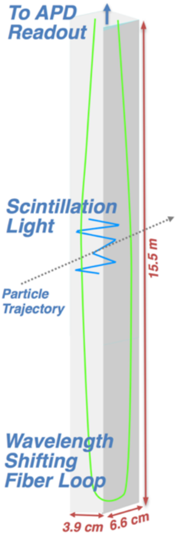


Near Detector ($L \sim 1\,\text{km}$, $M \sim 300\,\text{ton}$) measures original beam.
Far Detector ($L \sim 810\,\text{km}$, $M \sim 14\,\text{kton}$) measures oscillated beam.

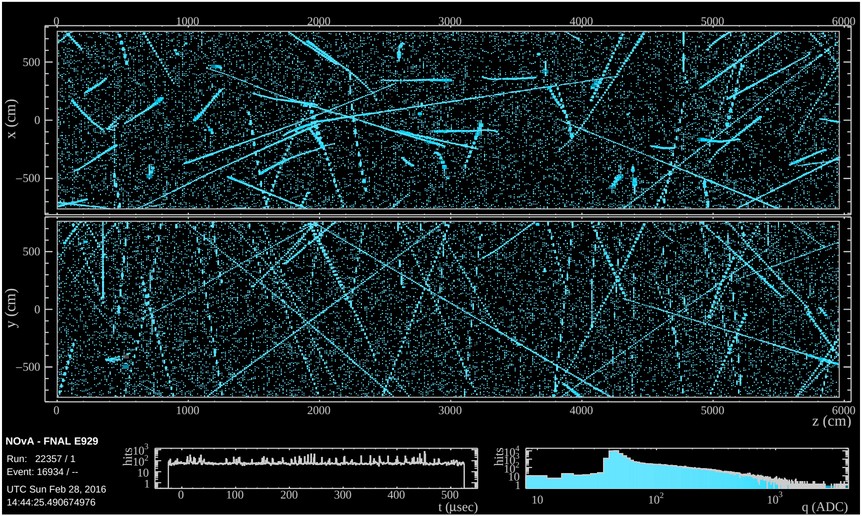
NOvA Detector Technology

- Basic unit of a detector is a long plastic tube with liquid scintillator (cell).
- Light is collected by an optical fiber and detected by an APD.
- Cells are combined into planes. Planes are stacked in alternating directions.





Sample of Activity in the NOvA Far Detector



550 μs window of Data

NOvA Physics

- NOvA performs two main analyses to constrain neutrino oscillation parameters:
  1. ν_μ Disappearance Analysis measuring $P_{\mu\to\mu}$
  2. ν_e Appearance Analysis measuring $P_{\mu\to e}$

  for neutrinos and antineutrinos.

- Sensitive to the atmospheric oscillation sector: Δm^2_{32}, θ_{23}, δ_{CP}.

- NOvA could help resolve some unanswered qustions about neutrino physics:
  - Neutrino Mass Hierarchy question?
  - Whether $\theta_{23} = \pi/4$?
  - Whether CP symmetry is violated in the neutrino sector?

ν_μ Disappearance Analysis

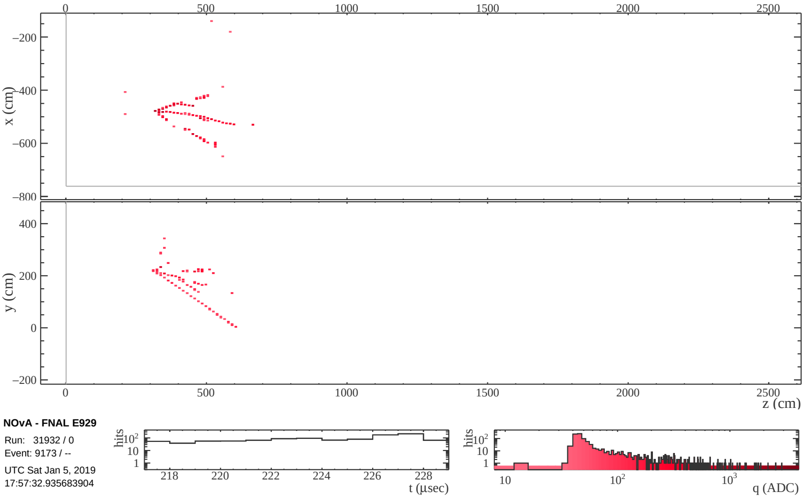
- ν_μ Disappearance Analysis is to estimate neutrino oscillation parameters $\{\Delta m^2_{32}, \theta_{23}\}$, by measuring survival probability of the ν_μ neutrinos at the Far Detector:

$$P_{\nu_\mu \to \nu_\mu}\left(E, L; \{\Delta m^2_{32}, \theta_{23}, ...\}\right)$$

- In order to make inferences about neutrino oscillation parameters $\{\Delta m^2_{32}, \theta_{23}\}$, we need to identify ν_μ neutrinos and estimate their energies.

- The only reliable way to identify ν_μ is when it interacts via the Charged Current interaction with the detector: $\nu_\mu \to \mu + \text{Had}$

Example of ν_μ Charged Current Event



Example of ν_μ CC event: $\nu_\mu \to \mu + \text{Had}$

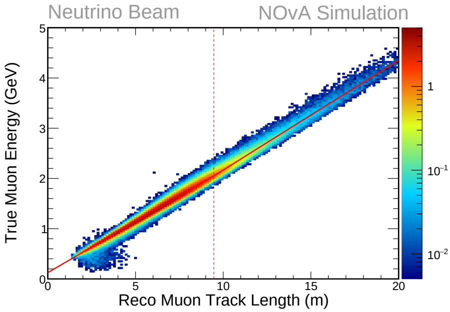
The Standard ν_μ Energy Estimator

- The Standard Energy Estimator of ν_μ CC events ($\nu_\mu \to \mu + \text{Had}$) exploits domain knowledge.

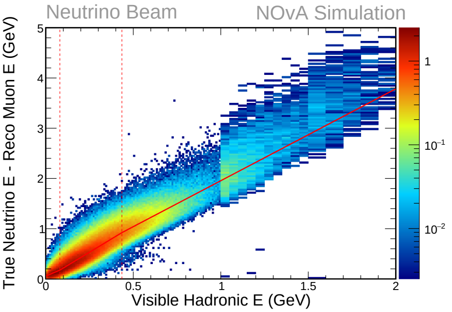
- It works in three steps:
  1. Identify μ track and estimate E_μ from its track length.
  2. Estimate E_{Had} from the calorimetric energy of its hits.
  3. $E_{\nu_\mu} = E_\mu + E_{\text{Had}}$

- It relies on the fact that μ tracks are relatively easy to identify, and that μ energy deposition rate dE/dx is well known.

The Standard ν_μ Energy Estimator, 2



(a) E_μ vs Muon Track Length



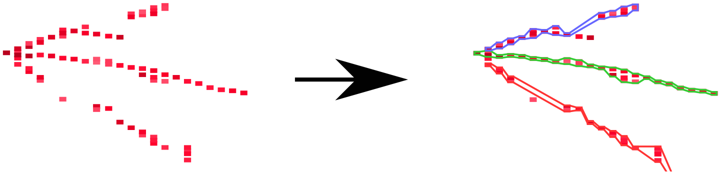
(b) E_{Had} vs Calorimetric Energy

Hadronic Energy component has large variance not explained by a total calorimetric energy

Can we estimate ν_μ energy better?

- The Standard ν_μ CC energy estimator has acceptable performance, since on average for selected events $2/3$ of E_{ν_μ} energy comes from E_μ and only $1/3$ comes from E_{Had}.
- How can we improve NOvA ν_μ CC energy estimator?
- At NOvA we can reconstruct clusters of hits (prongs), that correspond to individual particles at each event.

Particle Reconstruction



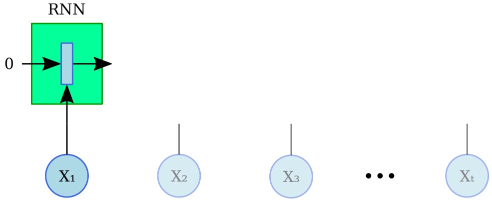
NOvA can reconstruct clusters of hits of individual particles:

- Find number of hits and calorimetric energies
- Estimate dimensions and directions
- Predict type of the particle
- Estimate energies and momenta of particles

RNN Energy Estimator

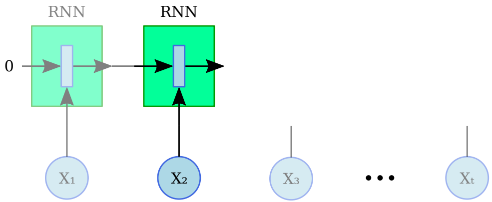
- We would like to use information from each particle as input to a new energy estimator.
- However, the number of particles (and prongs) varies between events.
- We needed a model that is capable of processing inputs of varying length.
- Recurrent Neural Networks are capable of handling inputs of varying lengths.

Recurrent Neural Network, 1



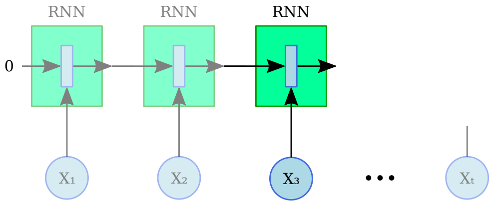
Recurrent Neural Network is a feed-forward neural network that is applied sequentially over inputs.

Recurrent Neural Network, 2



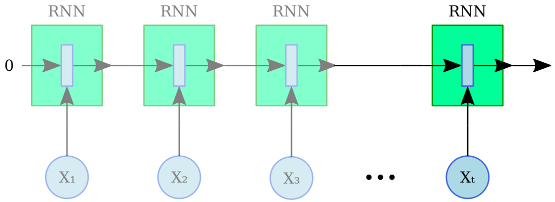
At each step network reads information from inputs and from the previous memory state, and outputs a new memory state.

Recurrent Neural Network, 3



At each step network reads information from inputs and from the previous memory state, and outputs a new memory state.

Recurrent Neural Network, 4



After all inputs have been processed, we extract output from the memory of the recurrent neural network.

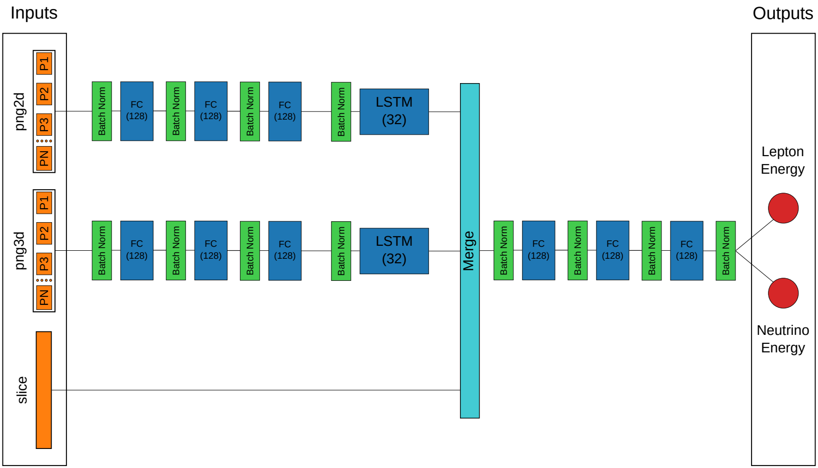
RNN Energy Estimator. Data Formats

- Rapid prototyping of Neural Networks is possible in Python and requires GPU enabled machines.

- NOvA dataset has size of about $\approx 1\,\text{TB}$ and cannot be easily accessed from Python nor transferred to a GPU cluster.

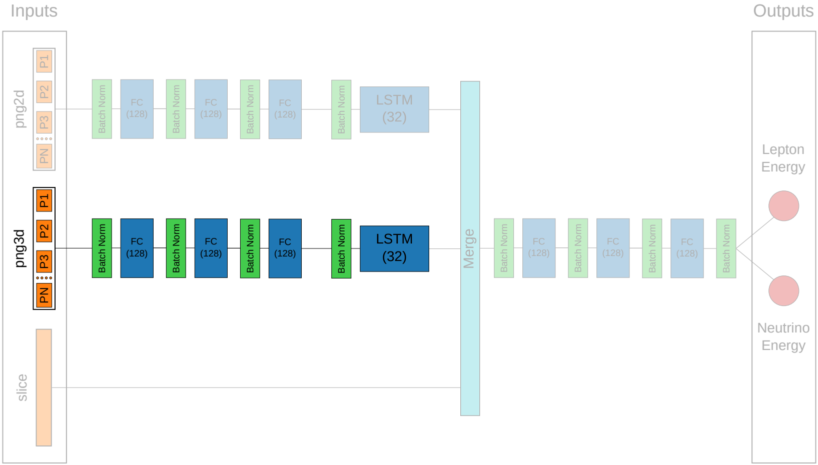
- I have designed an intermediate data format to extract relevant variables from NOvA ROOT files and transfer them to the GPU cluster, reducing dataset size down to $\approx 1\,\text{GB}$

Architecture of the Recurrent Energy Estimator, Overview



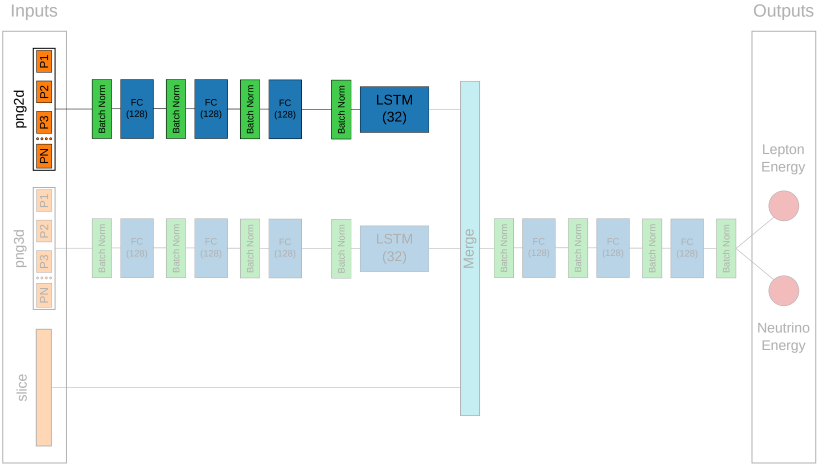
Long Short-Term Memory Cells are used to process fully reconstructed prongs (3D) and partially reconstructed prongs (2D)

Architecture of the Recurrent Energy Estimator, 3D Prong



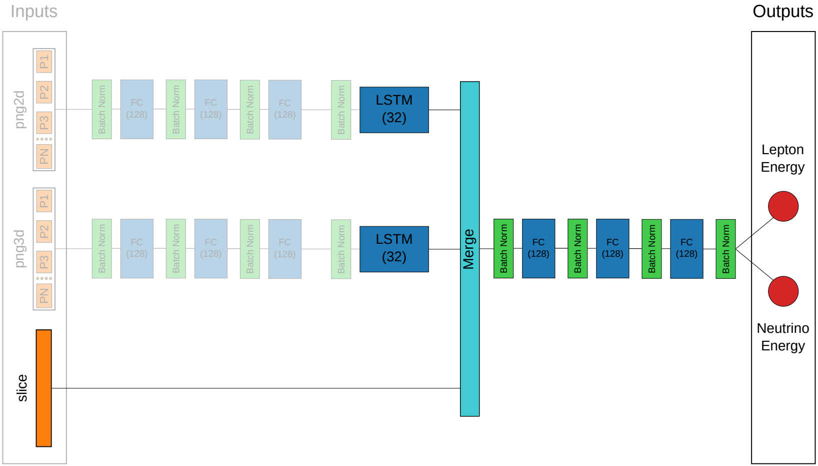
Information from fully reconstructed prongs (3D) is preprocessed through a set of Dense layers and fed to a LSTM Cell.

Architecture of the Recurrent Energy Estimator, 2D Prong



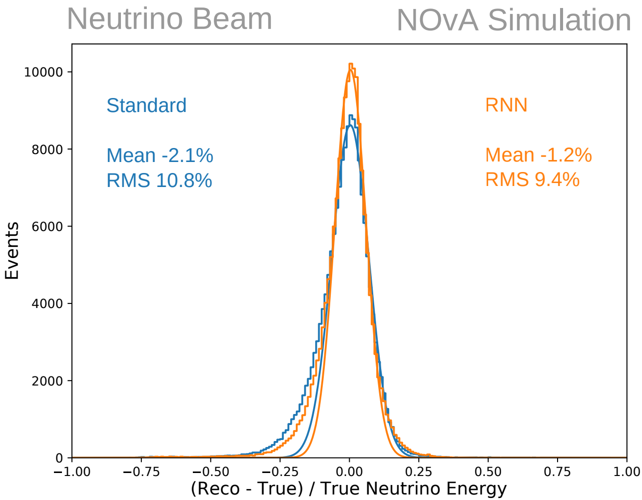
Information from partially reconstructed prongs (2D) is fed through another branch of Dense layers and LSTM Cell

Architecture of the Recurrent Energy Estimator, Output



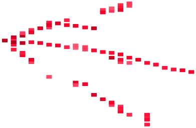
Outputs of LSTM Cells are combined with global information about event and used to predict μ and ν_μ energies.

Performance of the Recurrent Energy Estimator



RNN energy estimator is better than the standard in terms of RMS 9.4% vs 10.8%.

→ **CAT**



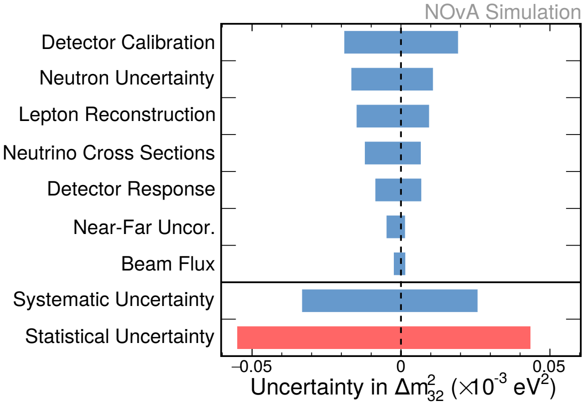
→ **?**

Humans cannot accurately identify event types, much less predict neutrino energies. We use Monte Carlo simulation to get labeled data.

Monte Carlo Simulation

- We use Monte Carlo simulation of neutrino interactions in order to train Machine/Deep Learning algorithms.
- Unfortunately, we do not have precise model of physical interactions, therefore results of this simulation are not fully accurate.
- We use systematic uncertainties in order to estimate errors of Monte Carlo simulation.

Systematic Uncertainties

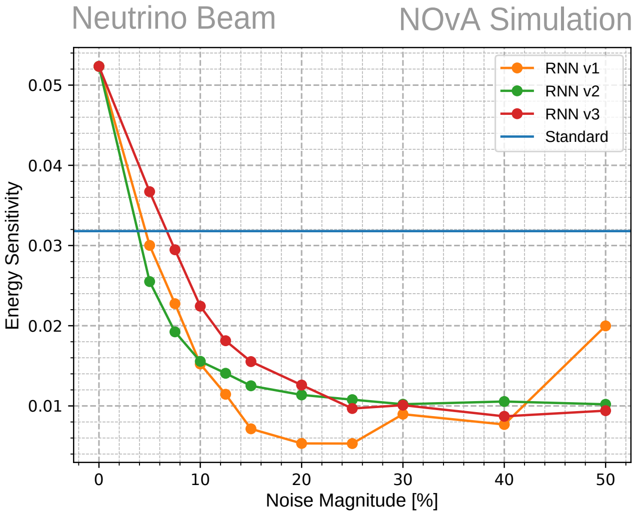


Precision of measurements of oscillation parameters is limited by systematic uncertainties.

Data Augmentation to Reduce Sensitivity to Systematic

- We would like to reduce sensitivity of the RNN energy estimator to the Calibration systematic uncertainty.
- It is possible to reduce sensitivity of an ML model to a systematic uncertainty of its inputs by adding random noise to the uncertain inputs.
- I have studied effects of addition of random noise in a way that emulates the effect of the Calibration systematic.

Sensitivity to the Calibration Systematic



RNN EE can be made 5 times less sensitive to the Calibration systematic than the Standard EE

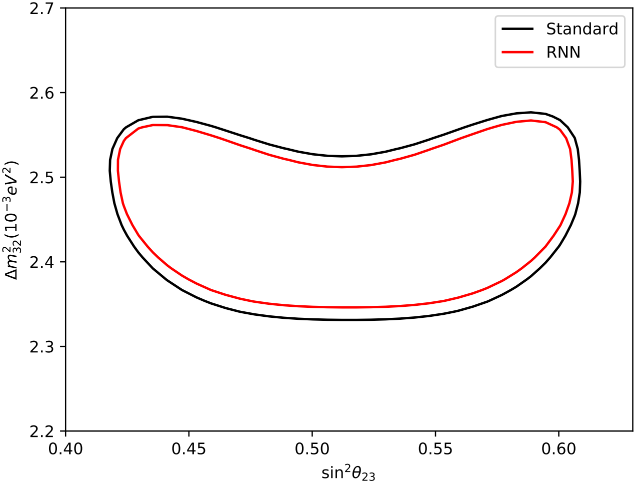
Effects of Using the RNN Energy Estimator

- New RNN energy estimator has 15% better energy reconstruction.

- New RNN energy estimator is 5 times less sensitive to the major systematic uncertainty at NOvA.

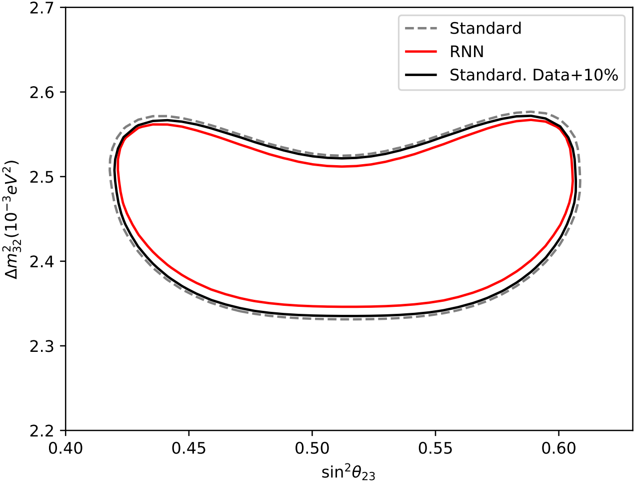
- **(Tentative Results)** Improvement due to usage of the RNN EE is equivalent to $10 - 50\%$ of additional data with the Standard EE.

NOvA Oscillation Parameter Contours



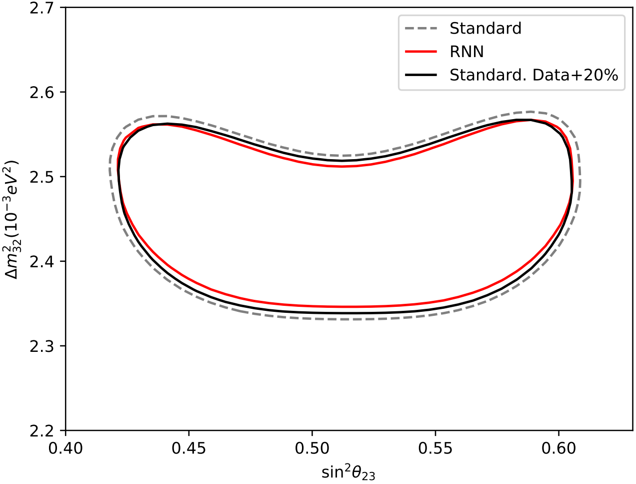
(Tentative Results) NOvA 1σ contours for Δm^2_{32} vs $\sin^2\theta_{23}$

NOvA Oscillation Parameter Contours, With 10% more Data



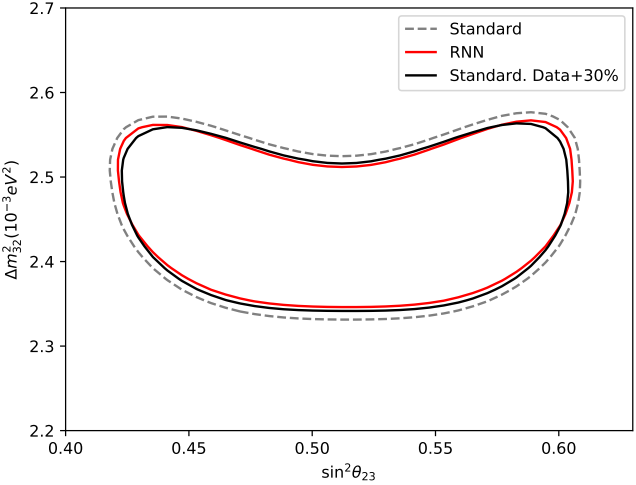
(Tentative Results) With of 10% of extra data the Standard EE performance does not match performance of the RNN EE.

NOvA Oscillation Parameter Contours, With 20% more Data



(Tentative Results) With of 20% of extra data the Standard EE performance matches RNN for $\sin^2\theta_{23}$

NOvA Oscillation Parameter Contours, With 30% more Data

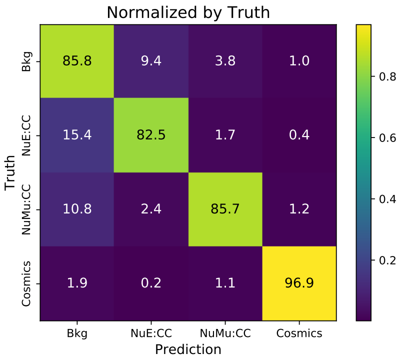


(Tentative Results) Even with 30% of extra data the Standard EE performance does not match RNN for Δm^2_{32}

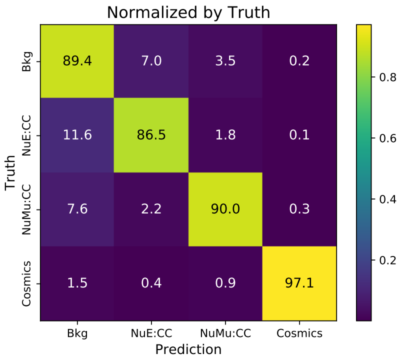
Usage of RNN Architecture for Event Classification

- NOvA relies on a CNN classifier in order to predict neutrino interaction type.
- However, CNNs are difficult to interpret from a physical point of view, and difficult to assess impact of systematic uncertainties on their output.
- NOvA needed an interpretable version of the event classifier to cross validate CNN classifier results.

Recall. RNN vs CNN



(a) RNN using only high-level information



(b) CNN using all available information

Usage of RNN Architecture for Event Classification, 2

- I have adapted the RNN energy estimator architecture to the task of event classification.
- The RNN event classifier has slightly lower performance (within 5%) compared to the CNN one, since it uses much less information as inputs.
- But the RNN classifier is easy to interpret and it is about 100 times faster to run.

Adoption of RNN Energy Estimator to Other Experiments

- I have developed an intermediate data format, data pipelines and the python package to train the RNN energy estimator, that are not specific to the NOvA experiment.
- They can be used to easily develop an RNN energy estimator for other experiments.
- Right now, I am porting the NOvA RNN EE to DUNE experiment, and it shows very promising results.

Conclusions

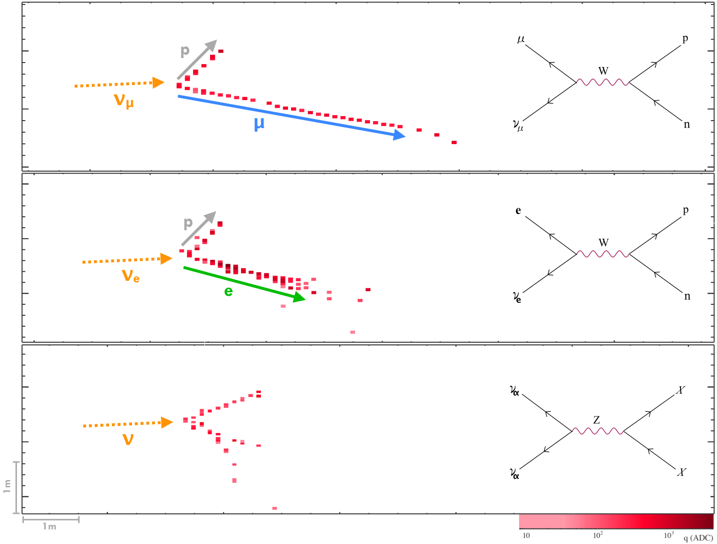
- I have developed an RNN energy estimator for the NOvA experiment, that has 15% better energy reconstruction and 5 times less sensitive to the major systematic uncertainty at NOvA.

- It may significantly improve performance of the NOvA experiment, pending further testing.

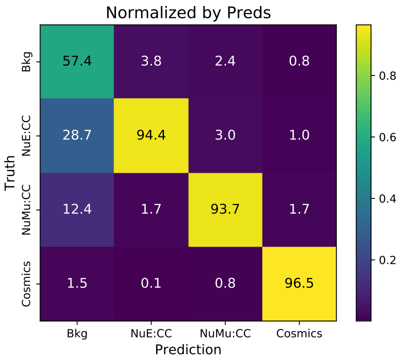
- The architecture of the RNN energy estimator can be easily adapted to the task of event classification, and ported to other experiments.

Backups

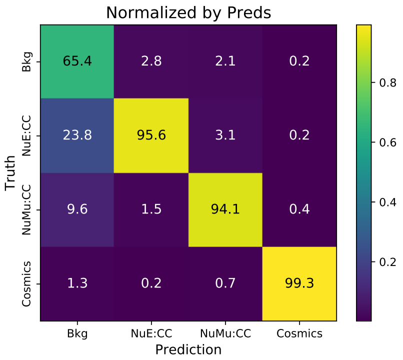
NOvA Event Topologies



Precision. SliceLID vs CVN

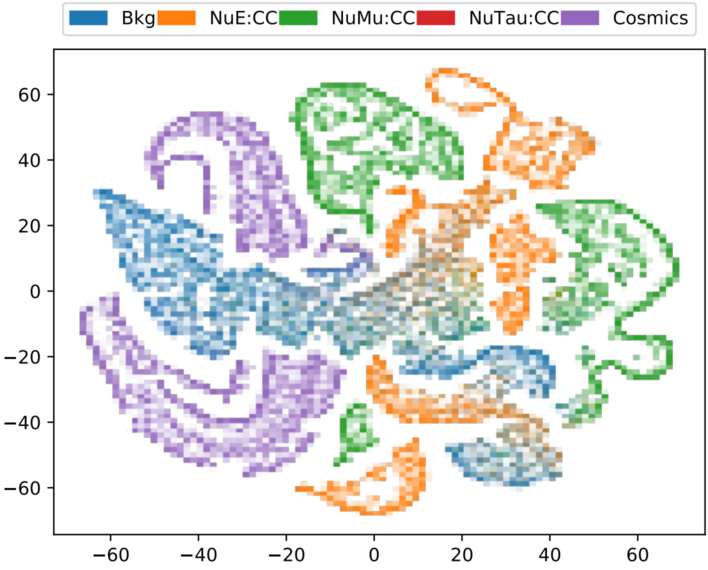


(a) RNN



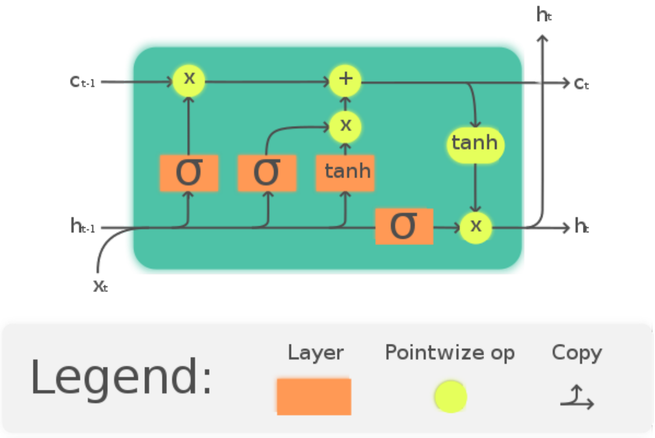
(b) CNN

t-SNE. SliceLID



t-SNE embedding of SliceLID outputs

LSTM Neural Cell



Source: `https://arxiv.org/abs/1808.05578`